



Visualizador de voz

Discriminación de fonemas

Introducción

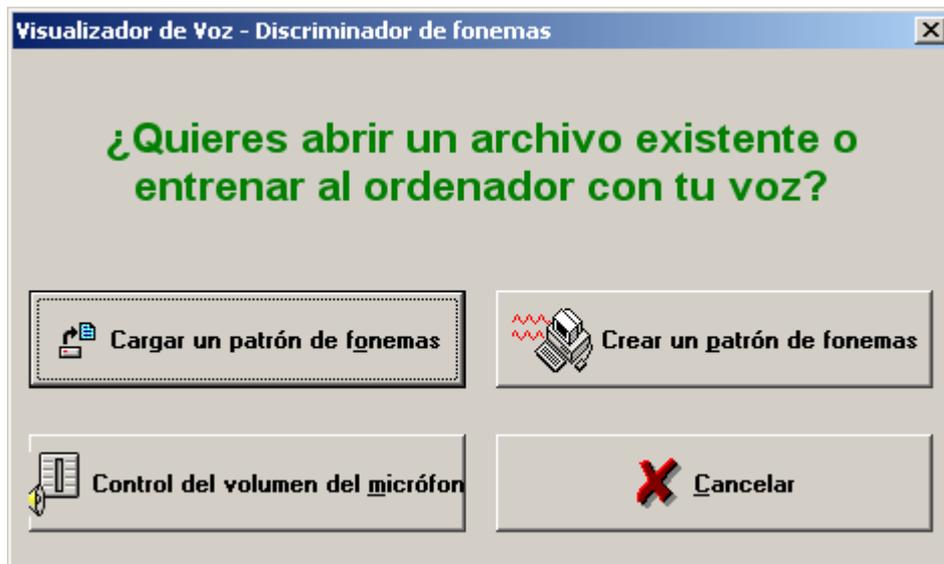
La intención de este programa es facilitar la discriminación de los fonemas a personas con deficiencias auditivas o con dificultades en el habla.

La idea para intentar conseguirlo es controlar sencillos juegos con la voz.

En principio el programa funciona comparando patrones de voz que el mismo programa permite crear.

Descripción de la fase de inicio del programa

Cuando encendemos el programa aparece después del cuadro de crédito el siguiente cuadro:



Aquí tenemos cuatro opciones:

Cargar un patrón de fonemas:

Nos permitirá recuperar un patrón de fonemas previamente entrenado i guardado en la opción **Guardar...** . Esta opción es la misma que la del menú **Abrir fichero**.

Crear un patrón de fonemas:

Esto nos permitirá construir patrones de fonemas. Esta opción es la misma que la del menú **Entrenamiento**.

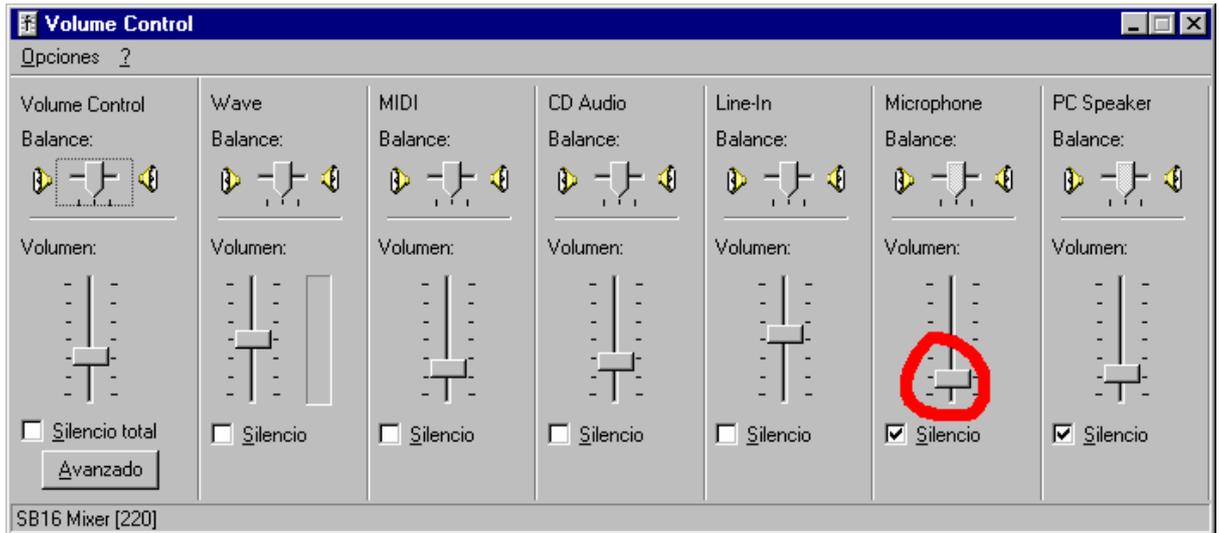
Control de volumen del micrófono:

Hace aparecer una ventana donde podremos ver el nivel de sonido de entrada cuando esté el sistema de captura de sonido en funcionamiento, o sea, cuando estemos **entrenando al ordenador** o cuando estemos haciendo el **reconocimiento**.



La barra de desplazamiento, la flecha, nos permite escoger el nivel mínimo de reconocimiento. ¿Qué quiere decir esto? Que los sonidos que no lleguen al nivel de la flecha no serán nunca reconocidos, y los que la pasen ocasionarán el proceso de reconocimiento. Este nivel se puede cambiar desplazando la flecha izquierda o derecha, dependiendo de si estamos en un ambiente silencioso o ruidoso, también puede interesar modificar esto si el nivel de detección de nuestro sistema micrófono-placa es muy o poco sensible.

Este cuadro también nos permite acceder al programa **Control del volumen**, programa que viene con Windows o con la tarjeta de sonido, y a partir de él subir o bajar el nivel del volumen de entrada de sonido.



Si no se ve ninguna señal hará falta subir el nivel del volumen del micrófono. Si por el contrario la señal se activa y llega al final muy fácilmente hará falta bajar el volumen.

Cuando aparece este cuadro se mantiene el cuadro anterior por si queremos hacer el **entrenamiento** o el **reconocimiento**.

Cancelar:

Si apretamos este botón no pasa nada, vamos directamente a la ventana principal del programa y podemos hacer las opciones que explicaremos a continuación:

Arrancar el programa con un parámetro:

Si arrancamos el programa con un parámetro, per ejemplo:

ReconeixementDeFonemes PepetAAEÈÉIOU.SonsTpmSR02

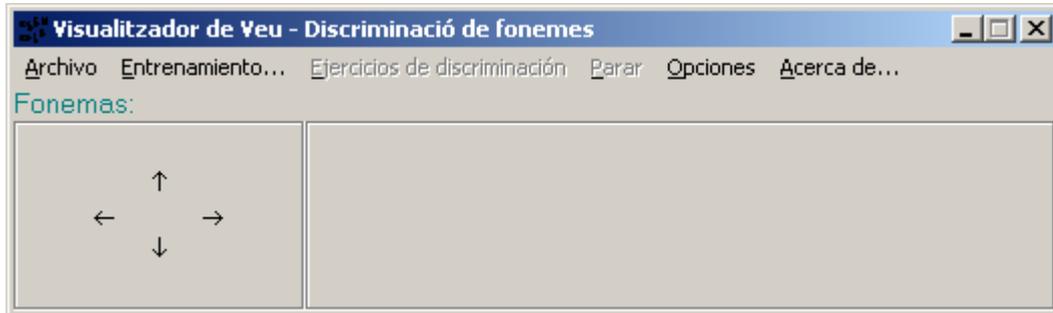
El programa cargará el fichero **PepetAAEÈÉIOU.SonsTpmSR02**, si contiene parámetros correctos de reconocimiento pasará ya la pantalla principal.

Recuerden que se puede hacer esto clicando el botón **Inicio** de Windows y, a continuación, clicando en **ejecutar programa** donde podemos escribir dicha línea de comandos.

También se puede hacer en el explorador de Windows clicando en el fichero y arrastrándolo encima del fichero.exe del programa.

Descripción de la ventana principal del programa

La ventana principal del programa tiene el siguiente aspecto:



- Vemos la barra del título. (Aquí se escribirá el fonema reconocido).
- Hay una barra de menús que describiremos más adelante.
- Debajo de la barra de menús hay un espacio gris. Aquí se escribirán los fonemas que puede entender.
- Un cuadro con unas flechas. Aquí se escribirán los fonemas que equivalen a las teclas que controlarán los juegos.
- A su lado, un cuadro donde aparecerá en letra grande el fonema reconocido cuando se haga el reconocimiento.

A continuación pasamos a describir las opciones de los menús:

Menú Archivo

Hay cinco opciones, hay algunas que se habilitarán o deshabilitarán, se vuelven grises, según el estado del programa.

Abrir... permite cargar un fichero de patrones que previamente hayamos guardado con este programa. Actúa de la misma manera que lo hemos explicado antes. La extensión por defecto de este programa es: **SonsTpmSR02**, tal y como siempre se recomienda, no es bueno cambiarlo. El programa, automáticamente, al iniciarse se asocia con esta extensión. Si hubiese otro programa con esta misma extensión, cosa bastante improbable, excepto que tengamos dos o más copias de este programa en nuestro disco duro, nos preguntará si se lo queremos asociar.

Guardar... permite guardar los datos de reconocimiento previamente recogidos en la opción. **Entrenamiento** en un fichero del disco. Tal y como hemos dicho antes la extensión por defecto es: **SonsTpmSR02** y no es conveniente cambiarla.

Abrir script... Se explica dicha opción más adelante en este manual.

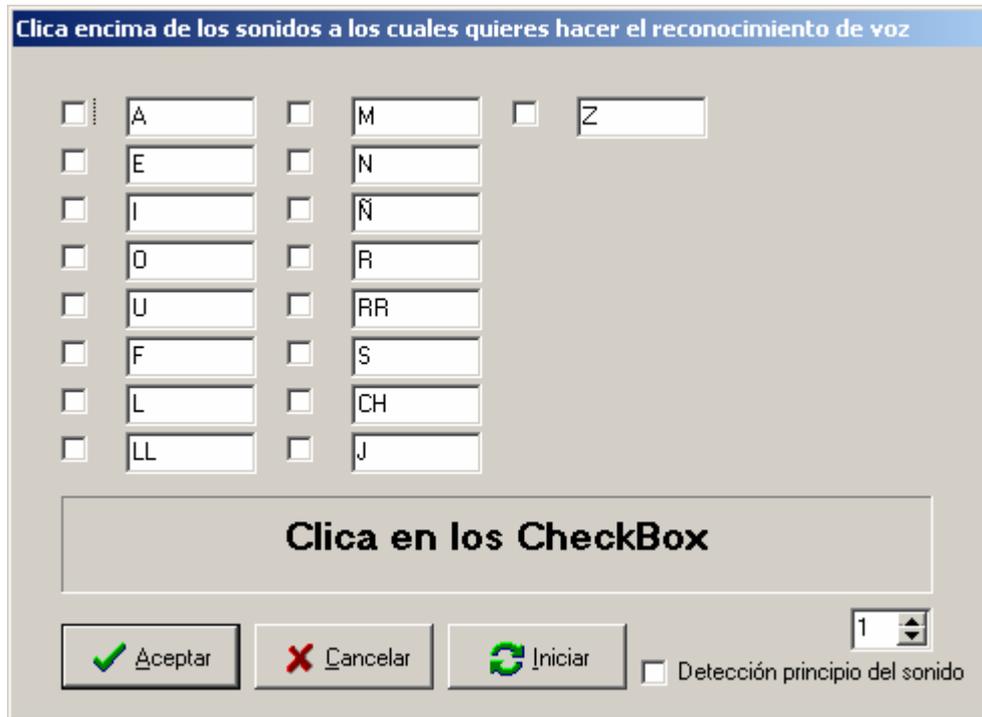
Ver el nivel del sonido... da acceso al mismo cuadro sobre el nivel del volumen del sonido que hemos explicado previamente.

Salir, permite abandonar el programa.

Menú **E**ntrenamiento...

Es la opción que permite que el ordenador aprenda a reconocer nuestra voz, o la de quien sea. Es la misma de la que hemos hablado antes en el cuadro inicial, pero que no hemos explicado en profundidad, y aquí pasamos a hacerlo. Empezamos a tratar las particularidades del programa ya que, hasta ahora, lo que hemos hablado, o cosas parecidas, acostumbra a encontrarse, en la mayoría de programas.

En primer lugar aparece el siguiente cuadro:



Aquí basta clicar en los sonidos a los que queremos crear el patrón de voz. Como mínimo hace falta clicar a dos sonidos.

Está la opción de cambiar los sonidos. En los cuadros de edición es donde están escritos los sonidos y es donde podemos poner los sonidos, sílabas, o las palabras que queramos. También, podemos poner el tiempo de duración del reconocimiento cambiando el valor del cuadro donde hay un 1, que significa 1 décima de segundo (como máximo 5 décimas de segundo). Si lo que queremos es que reconozca sílabas hará falta aumentar este número. También puede ser bueno clicar en el *checkbox* **Detección de principio de sonido** cuando queramos detectar sílabas o palabras. En

general para sonidos es mejor que esta opción esté desactivada y para sílabas y palabras es mejor que esté activada. La experiencia del usuario le indicará como modificar estos parámetros para que el reconocimiento sea mejor.

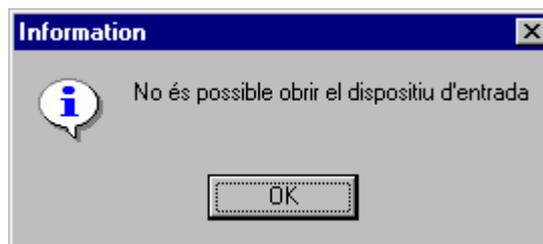
El botón **iniciar** hace que los nombres de los sonidos vuelvan a ser los que veis en esta figura.

Supongamos que hayamos clicado sobre todas las vocales: A, E, I, O, U, y clicamos en RETURN, nos aparecer la siguiente ventana:



Hará falta acercarnos al micrófono y decir el sonido *A* durante muy poco tiempo, sólo un par o tres décimas de segundo. Cuando la haya detectado desaparecerá del cuadro. Si desaparece demasiado deprisa, o sea, antes de que no hayamos dicho nada quizás sea porque tengamos el micrófono demasiado sensible. Recuerden lo que hemos explicado antes de utilizar la opción de control de volumen del micrófono. Al principio sería una buena costumbre tener esta ventana siempre abierta para ver si gritamos demasiado fuerte o flojo o tenemos el valor del micrófono demasiado arriba o abajo. Si por el contrario por más que gritamos no pasa nada, o sea, no desaparece este cuadro, puede ser debido a las causas antes mencionadas o a que el micrófono no esté correctamente enchufado a la tarjeta de sonido o el nivel del micrófono está demasiado bajo. En este último caso vayan a la opción **control del volumen**.

En el caso de que no estuviese la tarjeta instalada, que estuviese mal instalada, o que hubiese otro programa haciendo servir la entrada de voz, como puede ser la **grabadora de Windows**, antes de que saliese el anterior cuadro os saldría el siguiente:



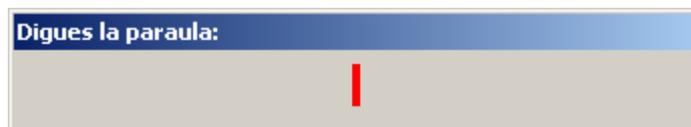
En este caso hará falta que os enteréis de cual de los problemas mencionados es el que falla y corregirlo. Hasta que no esté la tarjeta y el micro correctamente instalados, y hasta que no haya ningún programa que la haga servir, el programa no podrá funcionar (En windows XP parece que ya no es problema esto, ya que permite tener varios programas que capturan sonido a la vez). Si el ordenador tuviese más de una tarjeta no sé mucho qué puede pasar, este programa hace servir la que el ordenador ha

configurado como primera. Pero supongo que no se darán muchos casos de que esto pase.

Suponiendo que no haya habido esta dificultad y que haya aparecido la palabra a reconocer y el ordenador la haya detectado correctamente a continuación aparecerá un nuevo cuadro:



En este nuevo cuadro hacemos exactamente lo que acabamos de decir para entrenar dicho sonido. Si funciona, a continuación aparecerán los siguientes cuadros:

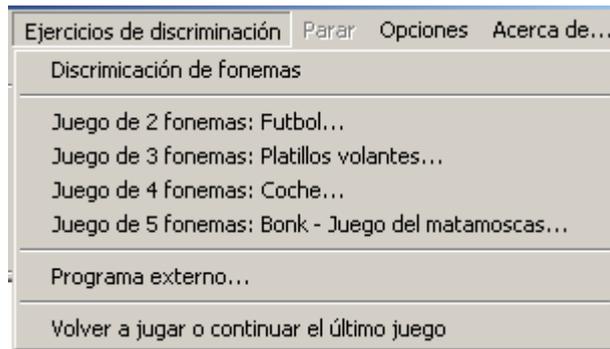


Si todo ha ido bien veréis como la ventana principal ha cambiado de aspecto y la opción del menú **Ejercicio de discriminación** habrá dejado de ser gris y será de color negro, que quiere decir que ahora ya está activa.

Cuando no estamos realizando el reconocimiento veréis que el menú **Guardar como...** está activado; esta será la opción que haremos servir si los patrones de reconocimiento son suficientemente buenos y queremos guardarlos para otra ocasión.

Menú Ejercicios de discriminación

Si lo activamos aparecerán 7 submenús:



Discriminación de fonemas:

Si lo activamos empieza el reconocimiento, veréis que los menús **Entrenamiento** y **Ejercicios de discriminación** se desactivan y se activa el menú **parar**. Si acercamos el micrófono a la boca y decimos sonidos parecidos a los fonemas que hay en el patrón que hemos grabado o recuperado del disco, veremos lo más parecido a lo indicado con letra grande en el cuadro de la derecha, que también aparece en la barra de título. Este es el ejercicio más simple porque el *feedback* sólo nos indica el fonema reconocido.

A continuación vienen 4 submenús donde se activan los juegos de discriminación de fonemas.

Estas cuatro opciones son equivalentes, la única diferencia es el número de fonemas a discriminar. Explicaremos la opción: **Juego de 3 fonemas: Platos volantes**.

Al clicar aparece el siguiente cuadro, que es muy parecido a uno anterior:

Escoge el sonido para la opción: Izquierda

<input checked="" type="checkbox"/> A	<input type="checkbox"/> M	<input type="checkbox"/> Z
<input checked="" type="checkbox"/> E	<input type="checkbox"/> N	
<input checked="" type="checkbox"/> I	<input type="checkbox"/> Ñ	
<input checked="" type="checkbox"/> O	<input type="checkbox"/> R	
<input checked="" type="checkbox"/> U	<input type="checkbox"/> RR	
<input type="checkbox"/> F	<input type="checkbox"/> S	
<input type="checkbox"/> L	<input type="checkbox"/> CH	
<input type="checkbox"/> LL	<input type="checkbox"/> J	

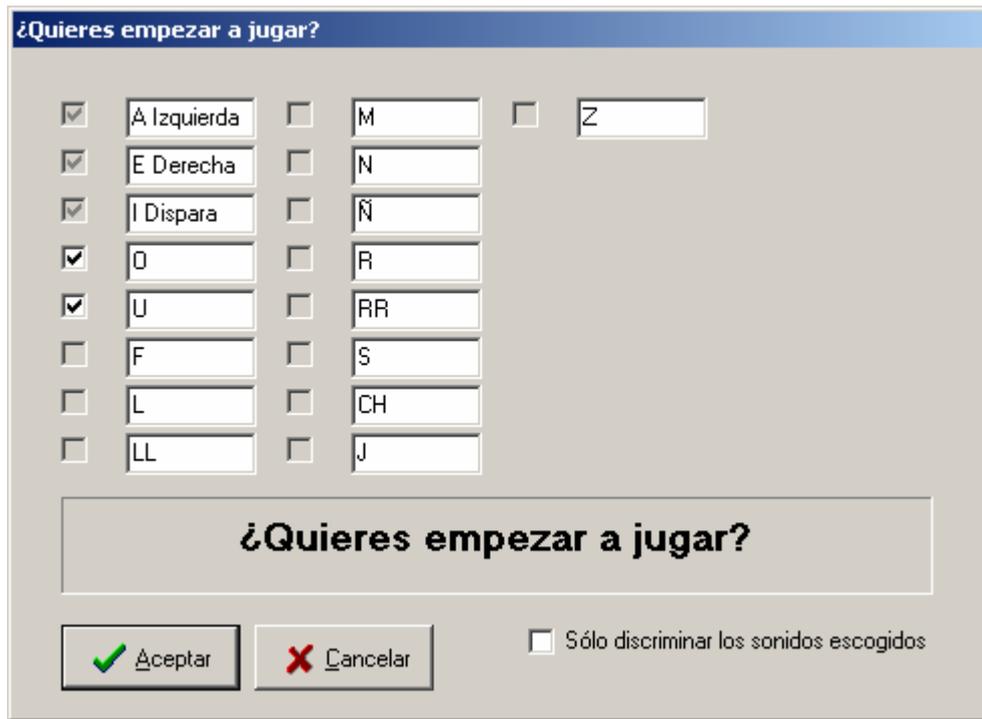
Clica en el CheckBox para: Izquierda

Aceptar

Cancelar

Sólo discriminar los sonidos escogidos

Vemos que sólo aparecen activadas las casillas de los fonemas que tenemos en el patrón. Aquí hará falta escoger tantos fonemas como teclas necesite el juego. En nuestro caso 3 teclas. Necesitaremos clicar consecutivamente 3 casillas para cada una de las teclas. O sea, primero para la tecla izquierda, después para la tecla derecha y finalmente para la tecla disparar, que son las 3 teclas que necesita el juego **platillos voladores**. Cuando hayamos clicado en los 3 fonemas aparecerá el mismo cuadro con el siguiente aspecto:



The dialog box has a title bar that says "¿Quieres empezar a jugar?". It contains a grid of checkboxes for selecting phonemes. The first three rows have their checkboxes checked: "A Izquierda", "E Derecha", and "I Dispara". The other checkboxes are unchecked. Below the grid is a large button with the text "¿Quieres empezar a jugar?". At the bottom left are two buttons: "Aceptar" (with a green checkmark) and "Cancelar" (with a red X). At the bottom right is a checkbox labeled "Sólo discriminar los sonidos escogidos".

<input checked="" type="checkbox"/>	A Izquierda	<input type="checkbox"/>	M	<input type="checkbox"/>	Z
<input checked="" type="checkbox"/>	E Derecha	<input type="checkbox"/>	N		
<input checked="" type="checkbox"/>	I Dispara	<input type="checkbox"/>	Ñ		
<input checked="" type="checkbox"/>	O	<input type="checkbox"/>	R		
<input checked="" type="checkbox"/>	U	<input type="checkbox"/>	RR		
<input type="checkbox"/>	F	<input type="checkbox"/>	S		
<input type="checkbox"/>	L	<input type="checkbox"/>	CH		
<input type="checkbox"/>	LL	<input type="checkbox"/>	J		

Vemos que se ha activado la tecla **Aceptar**. Si lo hacemos aparecerá el juego correspondiente, en este caso **platillos voladores**, y se activará el reconocimiento de voz y ya podremos controlar el juego con la voz. En este caso si decimos **A** veréis como el cohete va a la izquierda, si decís **E** el cohete irá a la derecha y si decís **I** el cohete disparará. De hecho al principio del juego aparece la primera pantalla donde se explica el funcionamiento del programa. Para empezar tenéis que disparar, o sea decir **I**.

Se puede parar el juego y el reconocimiento de voz en cualquier momento clicando el menú **parar**.

Si paráis el juego podéis reanudarlo clicando la opción del menú **Volver a jugar o continuar al último juego**.

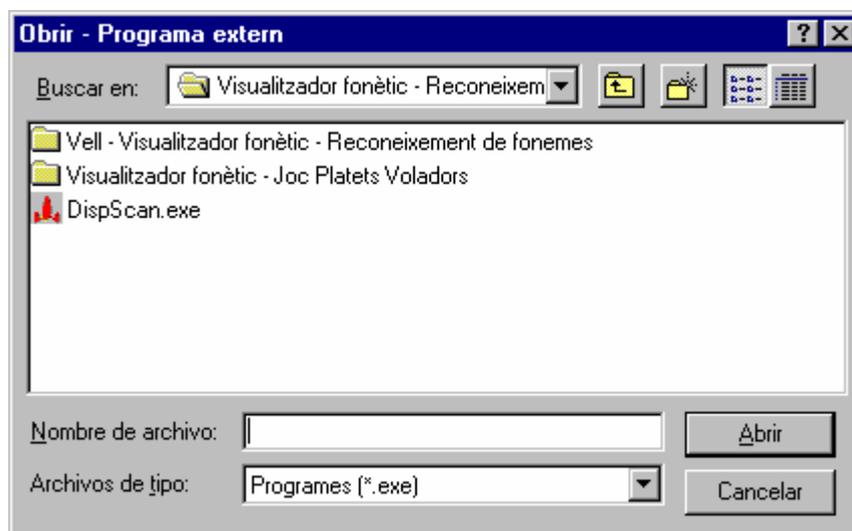
Nota 1: ChekBox **Sólo discriminar los sonidos escogidos**. Esta es una opción importante. Si nuestro patrón de sonidos contiene 5 fonemas y el juego necesita 3, por defecto el programa reconoce los 5 sonidos, o sea, que si decimos un sonido parecido a uno de los que no controla el juego se reconocerá éste y se indicará en la ventana, pero

no hará ningún efecto en el juego. Si se activa esta opción entonces sólo se comparará con los 3 únicos sonidos que controla el programa.

Nota 2: Para evitar problemas con la localización de los archivos necesarios para ejecutar los juegos, copiad el archivo, **cncs32.dll**, en el directorio, **windows/system**.

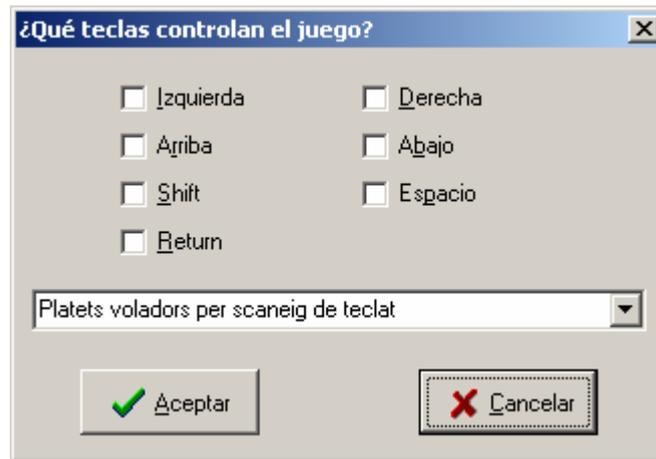
SubMenú Ejercicios de discriminación - Programa externo...

Esta opción es bastante importante para dedicarle un capítulo aparte. Está pensada para ampliar al máximo las posibilidades del programa. Si la clicamos aparecerá un cuadro estándar para abrir ficheros donde podremos escoger un programa para ejecutar.



Este programa puede ser en principio cualquiera. No tiene porque tener ninguna relación con el **projecte Fressa**. Si escogemos un programa y clicamos **Abrir** (el lenguaje depende del Windows que tengáis) se ejecutará el programa y al cabo de 5 segundos tiene que aparecer el siguiente cuadro. Si no apareciese, clicad en la ventana principal del programa **ReconocimientoDeFonemas**:

Si en el cuadro anterior clicamos **Cancelar** podremos utilizar otro juego que esté en pantalla escogiéndolo en el cuadro que ha aparecido.



Aquí hará falta clicar encima de las casillas que indican las teclas que controlan vuestro programa. Estas teclas están porque son las que creemos que controlan a la mayoría de juegos. Si alguno quiere hacer un juego controlado por voz haced servir estas teclas. Si alguien conoce un juego que quiere controlar con fonemas y no hace servir estas teclas hacéndonoslo saber.

Detalle importante, para que las teclas se envíen al programa correcto hace falta conocer el nombre de la ventana Windows que contiene el programa. En principio si todo funciona como tendría que funcionar, tiene que ser la que aparece en el cuadro de selección. **Platets voladors per escaneig de teclat**. Si no fuese esta la ventana haced doble clic en este cuadro de selección, se actualizan los títulos de ventana que en este momento se están ejecutando y después clicando en la flecha se desplegarán los títulos de las ventanas disponibles, clicad en la de vuestro juego o programa y ya podréis apretar de acuerdo y a continuación sucederá lo mismo que pasaba en los juegos del programa. O sea, saldrá el cuadro donde podéis asociar cada tecla con un fonema. A continuación se activará el reconocimiento y en principio podréis jugar al juego con reconocimiento de fonemas como si se tratase de uno de los juegos del programa. Hay que hacer notar que los juegos del programa son excluyentes, o sea, cuando jugamos a un juego el programa descarga los otros juegos que hay, pero no lo hace con los programas externos. Para descargar (salir) de los programas externos lo tendréis que hacer de la forma que lo haga el programa, si es una aplicación Windows normal, con Alt + F4 acostumbra a funcionar. Procurad descargar los programas cada vez que empiece un juego nuevo, sino no sabréis a que programa enviáis las teclas.

Menú Opciones

Si lo activamos aparecen 5 submenús:

Lenguaje. Que podemos hacer que el programa funcione en castellano o catalán.

Ver las funciones de onda. Sirve para detectar problemas de reconocimiento. Explicamos esta opción más abajo.

Sacar los juegos de la pantalla. Si queremos sacar los juegos del programa que hemos ejecutado previamente. También se puede hacer manualmente saliendo de ellos.

Intercambiar los botones de la rata. Esta opción sólo sirve por si se juega al juego de 5 sonidos, **bonk, juego del matamoscas**. Este juego no se controla por teclas, sino con la rata, se cazan moscas clicando. Si sois zurdos y hacéis servir el *mouse* con la mano izquierda es posible que tengáis los botones intercambiados. Si fuese así no cazaríais moscas nunca porque de hecho clicaríais con el otro botón. Con esta opción podéis intercambiar los clics que enviáis al programa **bonk**. En windows XP en principio no es necesario utilizar esta opción.

Scripts

Se puede crear un fichero script para facilitar la ejecución de los juegos controlados por voz. Se han de realizar de forma manual mediante un editor de textos como el, **notepad.exe**. En el mismo se ha de definir el fichero donde se guardan los patrones de voz, el nombre del ejecutable del juego y los sonidos asociados a las distintas teclas que controlan el juego. La extensión que deben tener los archivos script es, ***.srstpm**. La ventaja es que ejecutando directamente el script se carga el archivo de patrones el juego y la configuración de sonidos que activan las teclas para controlar el sonido. Puede ser abierto dicho script directamente desde el administrador de archivos de windows o bien desde el menú, **abrir**, submenú, **abrir script...** Como ejemplo se adjunta los archivos:

Ordres ES DIS DRE.srstpm Archivo script de ejemplo.

OrdresESDREDIS.SonsTpmSR02 Archivo con los patrones de sonidos, 3, Es, Dre, Dis. Izquierda, derecha, dispara.

El contenido del archivo de script:

```
Projecte Fressa 2001  Reconeixement de fonemes
Fitxer amb el patró de so = ordresESDREDIS.SonsTpmSR02
Fitxer amb el joc = DispScan.exe
//Tecles
// A(1) AE(2) È(3) É(4) I(5) O(6) U(7) F(8) L(9) LL(10) M(11) N(12)
Ñ(13) R(14) RR(15) S(16) SS(17) X(18) J(19) Z(20)
Esquerra = 1
Dreta = 2
Amunt = 0
Avall = 0
Shift = 3
Espai = 0
Return = 0
```

La primera línea es para que el programa *ReconeixementDeFonemes* reconozca el script como suyo. Siempre debe estar escrita así.

La segunda se escribe el nombre del fichero con el patrón de sonido, (*Fitxer amb el patró de so =*) en este caso el nombre de dicho fichero es, *ordresESDREDIS.SonsTpmSR02*.

Las dos líneas siguientes no sirven para nada. Cuando una línea va precedida de // el programa las ignora.

Las líneas a continuación indican que sonido se asigna a cada tecla. En nuestro caso la tecla Izquierda al primer sonido, *ES*, la tecla Derecha al segundo sonido, *DRE*, y finalmente la tecla disparar al tercer sonido, *DIS*.

Se recuerda que las palabras:

```
Projecte Fressa 2001  Reconeixement de fonemes
Fitxer amb el patró de so =
Fitxer amb el joc =
Esquerra =
Dreta =
Amunt =
```

Avall =
 Shift =
 Espai =
 Return =

No se pueden cambiar ya que son la sintaxis que entiende el programa.

Se adjunta otro ejemplo

Vocals AE I È.SrStpm
 vocalsAAEÈÉIOU.SonsTpmSR02

Se recuerda que los patrones han sido grabados con mi voz y utilizando fonemas catalanes por lo que a lo mejor si se quiere probar, éstos no funcionan correctamente.

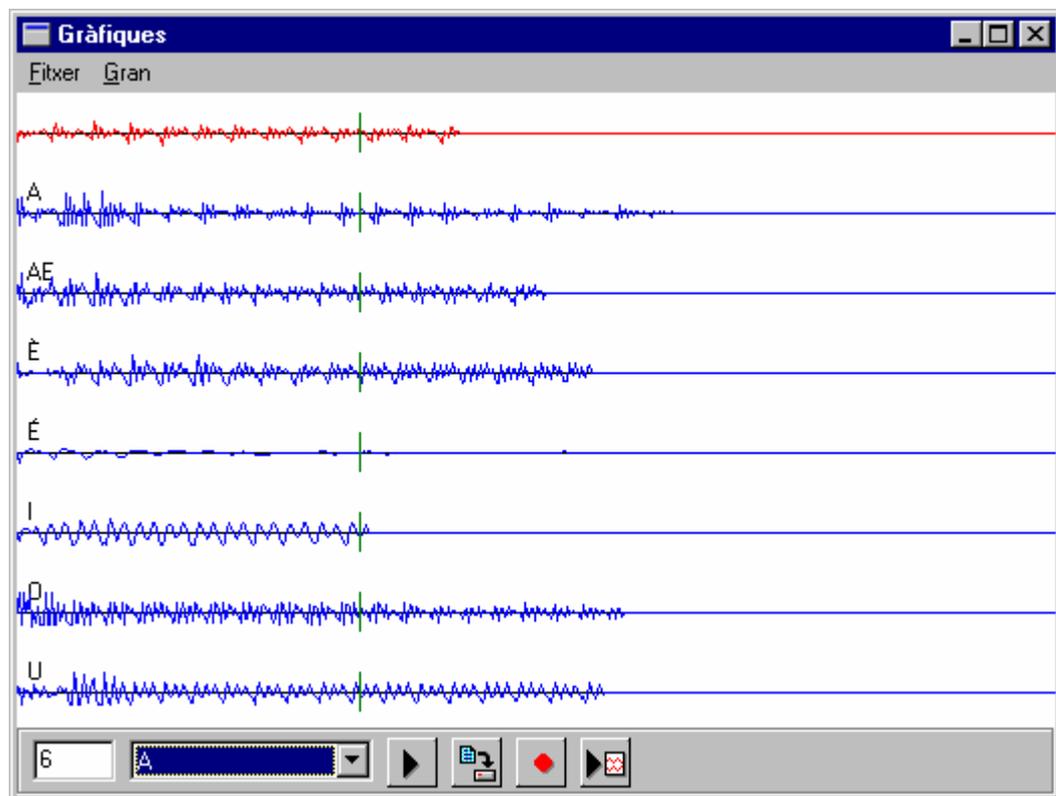
Qué pasa si detecta los sonidos mal

Dad un poco de confianza al programa, volved a repetir el entrenamiento y volvedlo a probar, al principio puede costar un poco, pero cuando funcione veréis que el resultado es espectacular.

¿Hay alguna manera de ver por qué falla el reconocimiento en el caso de que lo haga? Sí,

Menú **Opciones**, opción **Ver funciones de onda**.

Si clicáis esta opción os aparecerá la ventana siguiente:



Aquí tendréis el sonido que os ha grabado para hacer el reconocimiento. Si veis alguna gráfica de función que se vea muy poco quizás es porque no se ha grabado nada. Clicando en el último botón podremos escuchar todos los sonidos grabados, por si hay alguno con el volumen insuficiente o con otro sonido grabado que no es el que creemos. Con el primer botón podemos escuchar cada uno de los sonidos independientemente de los otros escogiéndolo en el cuadro de elección donde ahora aparece *AE-Izquierda*. Si algún sonido estuviese mal grabado podemos volver a hacer el entrenamiento para este último sonido clicando en el botón de color rojo que nos hará aparecer uno de los cuadros descritos anteriormente, y que en este caso sería:



Y podemos volver a grabar el sonido. El otro botón, el de la flecha señalando una especie de dibujo de un disco duro, sirve para guardar la grabación como fichero *.wav*. No es una opción que interese demasiado para el objetivo del programa..

Aprovecho para decir que si nos paseamos por encima de cualquier botón aparecerá un *tip* que nos dirá qué hace el botón.

El primer sonido, de color rojo, es al que le hacemos el reconocimiento y los azules son los patrones. Si tenemos escogido el primero veréis como el botón que tiene una redonda roja ésta le desaparece. Es porque el botón se vuelve inactivo. Este sonido no es un patrón.

Si alguna persona tiene cierta curiosidad y se pregunta si el reconocimiento se hace comparando estas funciones, os digo que no. Se hace haciendo unas transformaciones matemáticas a estas funciones.

Final

Visualizador de voz - Discriminador de fonemas es un programa que está continuamente en evolución. es posible que la versión que tengáis tenga más opciones que no estén recogidas en este manual o alguna cosa cambiada.

Visualizador de voz - Discriminador de fonemas se está desarrollando por Jordi Lagares Roset, profesor de matemáticas de el IES Santa Eugènia de Girona.

Visualizador de voz - Discriminador de fonemas es una primera aproximación al uso del reconocimiento de voz para ayudar a discriminar fonemas para personas con dificultades en el habla.

Visualizador de voz - Discriminador de fonemas es uno de los programas de los que forman el "**Projecte Fressa**" un conjunto de programas informáticos para ayudar en el mundo del discapacitado.

Aún y las limitaciones que pueda tener este programa, pensamos que puede ayudar a la escuela a cubrir un vacío de manera razonable reservando otras soluciones más costosas para equipos más especializados.

Para cualquier sugerencia y aportación, que será bien recibida, os podéis dirigir por correo electrónico a: jordi@lagares.org.

Podéis obtener más información, la versión más actualizada y otros programas del mismo autor en la dirección: www.lagares.org.

Animamos a los usuarios a visitar con cierta asiduidad dicha web ya que continuamente se van realizando modificaciones y actualizaciones de los programas del "**Projecte Fressa**".

Agradecimientos

El autor quiere expresar a *Joaquim Fonoll* del Programa de Informática Educativa y a las profesoras del **CRETAG** y de la escuela **Maçana** de Salt, el más sincero agradecimiento por el apoyo y sugerencias hechos a este programa. Así también quiero expresar mi agradecimiento a Mònica, mi esposa, por su inestimable ayuda en la confección de este manual.